Validation of PROMIS-57 in Kidney Transplant Recipients

Evan Tang¹, Aarushi Bansal¹, Sarah Cao¹, Nathaniel, Edwards¹, Madeline Li³, Marta Novak², Istvan Mucsi¹

¹ Multi-Organ Transplant Program, Department of Surgery, University Health Network

² Centre for Mental Health, University Health Network and Department of Psychiatry, University of Toronto

³ Department of Supportive Care, Princess Margaret Hospital

Background

Patient reported outcomes (PROs) are measures that are directly reported from patients. They allow measurement of patient factors not captured by traditional biomarkers. As such, they are a potential tool to improve the quality and outcome of care in kidney transplantation.

One challenge for PRO implementation is the lack of a generalized validated tool that is efficient and customizable. This need was recognized by the NIH resulting in the development of the Patient Reported Outcomes Measurement Information System (PROMIS).

PROMIS was developed and validated using item response theory which allows greater precision with less items (reducing response burden) and allows for flexibility and customization. In addition, the adjusted scale is normalized to the US population mean, enabling comparisons to be drawn. The goal of this study is the validate the PROMIS-57, a 57 question short form that measures depression, anxiety, fatigue, physical function, pain, sleep, and social functioning in the kidney transplant population.

Materials and Methods

A convenience sample of patients was recruited during transplant clinic and within the inpatient unit. Each participant completed the following questionnaires:

- PROMIS-57
- Kidney Disease Quality of Life Questionnaire
 - SF-12 Physical (PF) and Mental (MF) Composite
- Distress Assessment and Response Tool
 - Patient Health Questionnaire (PHQ9) – depression
 - Generalized Anxiety Disorder Scale (GAD7) – anxiety
 - Social Difficulties Inventory (SDI)
 - Edmonton symptom assessment system (ESAS)
- European Quality of Life Questionnaire (EQ-5D-5L)

Validation was explored using reliability and validity measures. Reliability measures were internal consistency (Cronbach’s α) and test-retest reliability (Intraclass Correlation Coefficient). Validity measures were structural validity (Principal component factor analysis), construct validity (Spearman’s Rho), and goodness of fit (Comparative fit index and Tucker-Lewis index).

Table 1. Population Demographics (n = 153)

Age ($\pm$ SD)	49.9 ($\pm$ 17)
Gender (%)	
Male	87 (57)
Female	66 (43)
Ethnicity (%)	
Caucasian	88 (62)
Black	19 (13)
Asian	33 (22)
Other/missing	13 (8)
Education (%)	
Less than high school	10 (7)
High School	40 (26)
Greater than high school	10 (65)
CCI > 2 (%)	55 (38)
CCI: Charlson Comorbidity Index	

Table 3. Construct Validity

Promis-57 Domain	Promis Mean (min,max)	Legacy Instrument	Legacy Mean (min,max)	Spearman’s Rho (95% CI)
Anxiety	48.9 (37.1,83.1)	GAD7	2.8 (0,21)	0.754 (0.647 to 0.832)
Physical Function	46.9 (20.9,59.7)	SF-12 PF	45.5 (0,100)	0.840 (0.771 to 0.889)
Depression	47.2 (38.2,81.3)	PHQ9	3.6 (0,27)	0.639 (0.495 to 0.749)
Fatigue	49.3 (33.1,77.8)	SF-12 PF	45.5 (0,100)	-0.603 (-0.686 to -0.541)
Pain	50.4 (50.7,77.0)	SF-12 PF	45.5 (0,100)	-0.683 (-0.775 to -0.562)
Sleep	47.8 (30.5,77.6)	PHQ Sleep	0.7 (0,3)	0.684 (0.555 to 0.781)
Social Functioning	52.6 (25.9,65.4)	SDI Score	4.6 (0,44)	0.754 (0.647 to 0.832)

|Rho| > 0.8: very strong, 0.6–0.8: Strong Correlation

Table 2. Reliability measures and Goodness of fit

PROMIS-57 Domains	Cronbach’s alpha (α)	ICC	CFI
Anxiety	0.96	0.73	0.969
Physical Function	0.96	0.89	0.925
Depression	0.96	0.84	0.967
Fatigue	0.97	0.76	0.974
Pain	0.99	0.75	0.958
Sleep	0.93	0.83	0.880
Social Functioning	0.97	0.77	0.900

ICC: Intraclass Correlation Coefficient
CFI: Comparative Fitness Index
 $\alpha > 0.9$: excellent internal consistency
ICC > 0.9: excellent reliability, 0.75–0.9: good reliability
CFI > 0.9: good fit, 0.8–0.9: permissible fit

Discussion and Conclusion

PROMIS-57 shows good validity and reliability in the kidney transplant population. A ceiling and floor effect for some of the PROMIS domains. One explanation is social desirability bias which may skew the responses. Alternatively, the range of domains may exceed the measurable range of PROMIS-57. In addition, convenience sampling potentially selects for higher-functioning individuals.

Future directions include the validation of computer adaptive testing; providing further reduction in responder burden and increasing the measurable range. Exploring the clinical significance of PROs in relation to traditional outcomes such as mortality and graft failure can serve to inform clinical practice.